Available online at www.sciencedirect.com

ScienceDirect

Lingua 136 (2013) 16–37

Lingua

www.elsevier.com/locate/lingua

Coreference, lexical givenness and prosody in German

Stefan Baumann [a,*], Arndt Riester [b]

[a] IfL Phonetik, University of Cologne, Herbert-Lewin-Strasse 6, 50931 Köln, Germany
[b] IMS, University of Stuttgart, Pfaffenwaldring 5b, 70569 Stuttgart, Germany

Received 26 May 2011; received in revised form 21 February 2013; accepted 18 July 2013
Available online 4 September 2013

Abstract

In this article we discuss some empirical results concerning the impact of different levels of information status (i.e. referents and words, respectively) on the prosodic realisation of referential expressions in annotated corpora of read and spontaneous speech. Both at the referential and at the lexical level not only given and new but also intermediate classes of givenness/novelty have to be distinguished. We provide a brief introduction to our two-dimensional RefLex annotation scheme and discuss its application to a number of examples from the theoretical literature which cannot be described satisfactorily by means of previous annotation schemes. From these examples we derive hypotheses on the relationship between information status and accent position as well as accent type. The hypotheses are generally confirmed for read speech showing a stepwise increase in prosodic prominence from given to new items, predominantly ordered according to the information status at the lexical level. The results of the relationship between prosody and information status are found to be less clear in spontaneous speech, probably due to the production of shorter intonation phrases reducing the variability of accentuation in marking different levels of givenness.

Keywords: Givenness; Information status; Annotation; Prosody; Coreference; Lexical relations

1. Introduction

This paper examines the question whether different types and degrees of givenness of referential expressions trigger different prosodic markings. We claim that an item's givenness may have two sources which are orthogonal to each other: it may not only stem from *coreference* with an antecedent but also from the *availability of a lexical unit* in the discourse. Consider the following two examples[1]:

(1) Yesterday, a friend of mine prepared a laSAgne for me.
  a. I found it hard to enJOY the tasteless stuff.
  b. It tasted much better than my OWN lasagne.

Both continuations involve deaccentuation on the sentence-final expression, which represents a deviation from the default realisation in an all-new setting. We claim that there are two different types of linguistic processes leading to the two cases of deaccentuation, which, however, have both been treated under an undifferentiated notion of *givenness*.

* Corresponding author. Tel.: +49 221 470 4259; fax: +49 221 470 5938.
E-mail addresses: stefan.baumann@uni-koeln.de (S. Baumann), arndt.riester@ims.uni-stuttgart.de (A. Riester).

[1] Capital letters indicate relevant (mostly nuclear) accents. The constituents in question (in most cases annotated expressions and their antecedents) are underlined.

0024-3841/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.lingua.2013.07.012

In Baumann and Riester (2012), we have developed a concise coding scheme for linguistic data, which enables the annotator to make a distinction between the cases in (1a) and (1b), as well as to reliably classify referential expressions in general. The motivation of the current paper is to subject this annotation scheme to empirical scrutiny. To this end, we annotate and investigate two types of German data, namely read episodes which are semantically controlled for, as well as spontaneous monologues. Most studies on the relationship between information structure and prosody examine controlled data only. However, we assume that this relationship may be different for various text types. The present study thus wants to verify this assumption by additionally looking at an uncontrolled, i.e. spontaneous, data set.

The paper is structured as follows: Section 2 defines givenness (or information status) as a nuanced cognitive notion, which – in intonation languages like German and English – is to a large extent expressed prosodically. It is shown that not only presence or absence of an accent but also the type of an accent as well as its position in an intonation phrase (i.e. prenuclear, nuclear, or postnuclear) serve as relevant markers of information status.

In Section 3, we propose a scheme for the annotation of a referring expression's information status, which is based on both lexico-semantic and morpho-syntactic information. In particular, we suggest that an item's information status should be investigated at two distinct levels, namely a *referential level* and a *lexical level*. The annotation scheme is therefore called the *RefLex scheme* (for a more detailed presentation of the scheme, including an in-depth discussion of the theoretical background, see Baumann and Riester, 2012). We claim that these two levels of information status have an impact on a nominal expression's degree of prosodic prominence. In Section 4, we will present hypotheses about the prosodic marking of nominal expressions, combining both levels of information status, which are tested on two German datasets (Section 5). Finally, in Section 6, we will draw some conclusions from the results. In sum, they show that each level has an incremental effect on the degree of prosodic prominence assigned to an expression. We find a general tendency of a stepwise increase in prosodic prominence from given to new items, which is particularly clear in read speech.

2. Givenness and prosody

A communicative situation consists (inter alia) of the sum of propositions the participants in a conversation are aware of and believe to be valid at the time of an utterance. This set of propositions is often referred to as *shared knowledge* (e.g. Clark and Haviland, 1977) or *common ground* (e.g. Stalnaker, 1974; Chafe, 1976; Krifka, 2007). Büring (2002) takes this notion as the basis for his definition of *informativity*:

> A sentence is informative with respect to a common ground if (and only if) the proposition expressed by the sentence is not entailed by the common ground.

For illustration, see example (2):

(2) A: What did you buy?
B: I bought a CAR.

Sentence (2B) is informative, since the argument *a car* of the proposition {B bought a car} has not been part of the knowledge stock of *both* interlocutors, adding the whole proposition to the common ground (although the open proposition {B bought X} was already included in the common ground).

This shows that conveying information is relational in yet another aspect: a speaker usually says something *new* with respect to what is already *given*. The denotata of individual sentence constituents, which are present in the minds of speakers and listeners, can be regarded to have cognitive states which might again be called given or new. While the propositional information {B buys X} is given by the question in (2A), the referent of X is new in (2).

The notion of ''referent'' usually applies to the *idea* or *mental representation* of people, objects or abstractions, not to their correlates in the external or ''real'' world. (They might not exist at all). Thus, what we are interested in is the (non-linguistic) image of a referent in the minds of speaker and listener and the linguistic expression denoting this image. In the following we will concentrate on the *prosodic marking* of referential expressions.

It has been shown in several studies on West Germanic languages that givenness triggers deaccentuation (which means that an element does not receive a pitch accent although it would be expected to get an accent in the ''unmarked'' case, cf. Ladd, 1980),[2] while newness triggers accentuation (see e.g. Halliday, 1967; Allerton, 1978; Cruttenden, 2006).

[2] Since we do not want to be biased by a prescriptive notion of where an accent *should* be placed, we actually use the term ''deaccentuation'' synonymously with ''lack of accentuation''.

The determiner phrase *a car* in (2B), for instance, is new and receives an accent. In the second sentence of example (3), on the other hand, the DP *the dog* is likely to get deaccented since it refers to the same dog as in the first sentence:

(3) On my way home, a dog barked at me. I was aFRAID of the dog.

The vast majority of cases of anaphoric reference are probably of the kind in (3) (including cases of pronominal anaphors), suggesting that the basic source of givenness and, in turn, deaccentuation, is *coreference*. In support of this view, there are many cases of deaccentuation in which the coreferring anaphor is not even expressed by the same term, as in (4) (from Büring, 2007:4; adapted from Ladd, 1980):

(4) A: Did you see Dr. Cremer to get your root canal?
B: Don't remind me. I'd like to STRANgle the butcher.

Here, the deaccentuation of *the butcher* clearly indicates that the item has to be interpreted as coreferential with *Dr. Cremer*. In contrast, an accent on *the butcher* would mark the expression as denoting a new – or, at least, a different – referent.

However, coreference is not the only source of deaccentuation. Often, the repetition of a lexical item prevents accentuation, even if the two items do *not* corefer. See example (5):

(5) On my way home, a dog barked at me. It made me think of ANna's dog.

The two instances of the lexical item *dog* do not refer to the same creature, in contrast to example (3) above. Nevertheless, although the two expressions are not coreferential they do at least denote the same concept, i.e. a dog. A distinction between a referential and a lexical level is also implied in Schwarzschild's (1999) notion of givenness, as well as in the system of cohesion within the framework of Systemic-Functional Linguistics (e.g. Halliday and Hasan, 1976; Halliday and Matthiessen, 2004).

Deaccentuation due to lexical relatedness does not only apply to repetitions of the same expression, but is also possible with certain ''bridging'' inferences (Clark, 1977), as in (6):

(6) Because of the broken drawers we had to get RID of the desk.

Here, the mention of *the broken drawers* suffices to render *the desk* no longer entirely new, since the latter expression subsumes (i.e. is a holonym of) the first.[3] Van Deemter (1994:21) calls this phenomenon *concept-givenness*. Note, however, that such subordinate-superordinate relations are asymmetrical: whereas the anaphoric holonym *desk* can be deaccented, as in (6), the anaphoric meronym *drawers* cannot, resulting in (7):

(7) We had to get rid of the desk because someone had broken the DRAWers.
but * We had to get rid of the desk because someone had BROKEN the drawers.

That is, meronyms like *drawers* in (7) would not be concept-given, but in some sense new, and thus have to be accented. To summarise the claims made by use of these constructed examples from the literature, we might say that deaccentuation may be caused by different types of givenness: (i) givenness of a discourse referent, (ii) givenness of a lexical item, and (iii) ''concept-givenness'' due to inferential relations between two entities. Conversely, other kinds of information status which lack any kind of relatedness lead to accentuation. Clearly, however, we would like to look beyond this introspective picture derived from the theoretical literature. In the discussion in Section 3, we will therefore sharpen our insights into annotating information status, and refine our terminology. In Sections 4 and 5 we will, finally, present empirical data which will either support or refute the assumptions just made, and will refine them with regard to more prosodic detail.

A model that allows for some gradience (or at least several steps) in the description of givenness is the ''cognitive''[4] approach proposed by Chafe (1976, 1994) and Lambrecht (1994). Chafe defines givenness in terms of the *activation cost* a speaker has to invest in order to transfer an idea from a previous state into an active state. He postulates an intermediate level between the poles *given* and *new*, referred to as *accessible* information.

[3] The same holds for all subordinate-superordinate relations, e.g. a hyponym (dog) followed by a hypernym (animal) (see also example (12) in Section 3 below).

[4] We use inverted commas because Chafe's approach is not based on neurocognitive evidence.

More recently, the activation cost model has received further support from neurolinguistic experiments analysing event-related brain potentials (ERPs) during reading comprehension tasks on German (Burkhardt, 2006, 2007; Burkhardt and Roehm, 2007). The studies investigated real-time sentence comprehension looking at which point inferential knowledge influences the cognitive processing of nominal expressions. It could be shown that two important ERP components, N400 and P600,[5] mark distinct processes relating to the given-new distinction. An enhanced N400 reflects difficulties during linking with the discourse model, and its amplitude increases with increased newness. The late positivity found in the reading studies was observed for costs from discourse updating and restructuring. In particular, Burkhardt (2006) supports the idea that accessible expressions (elicited via a bridging inference from an antecedent) take an intermediate position in terms of cognitive processing costs since they share properties with both given (showing an attenuated N400) and new information (showing an enhanced P600).

A recent ERP study on the *prosody* of inferentially accessible items in German (Schumacher and Baumann, 2010) revealed that the different *accent types* investigated evoked N400 modulations resulting in a statistically significant three-way distinction: deaccentuation led to a more pronounced N400 than a high pitch accent (notation in GToBI: H*; see Grice et al., 2005), which in turn was more pronounced than a high-falling or ''early peak'' accent (GToBI: H + L*). That is, H + L* proved to be the most adequate accent type for accessible information since it required the least processing effort (although this accent type takes a medial position in terms of perceived prominence; see e.g. Ladd and Morton, 1997).

This neurolinguistic study nicely confirmed that the presence or absence of a pitch accent is not the only correlate of an item's degree of activation or, respectively, its information status. Both accent *type* and *position* convey relevant meaning differences as well, roughly corresponding to the degree of prosodic prominence they evoke. As to accent type, it has been suggested that high and/or medial peak accents (H*) are used to mark new discourse referents, whereas downstepped (!H*) and/or early peak accents (H + L*/H + !H*) mark accessible ones, while low accents (L*) mark given items – if these are accented at all. This simplified classification seems to hold for English (Pierrehumbert and Hirschberg, 1990) and German (Kohler, 1991; Baumann, 2006; Baumann and Grice, 2006) and suggests that the pitch height of the accented syllable correlates with an item's degree of givenness, which is in line with Gussenhoven's (2004) *Effort Code*: a new item is attributed more importance for a speaker and is thus produced with higher pitch involving more articulatory effort.

The three-way accent type distinction measured in Schumacher and Baumann (2010) could also be confirmed in a prosodically more fine-grained auditory perception study on German by Röhr and Baumann (2011).[6] In an online experiment, subjects had to judge an item's degree of givenness (to be marked on a gradual scale between the poles ''known'' and ''new'') solely on the basis of its prosodic realisation, i.e. without any context. Results are presented in Fig. 1, indicating three significantly different groups of pitch accent type (and position). The most important criterion for the differentiation of nuclear accents seems to be the falling/low(er) component of the accents indicating increased givenness and the rising/high(er) main component of H* and !H* accents (indicating increased newness).

Based on these results we will compare four accent types in the tonal analysis of the data discussed in the present paper: H*, !H*, Early Peak (''EP''), and L* (disregarding deaccentuation here). That is, we conflate the two early peak categories proposed in GToBI (H + L*, H + !H*) into a single category, which has already been suggested by Grice et al. (2009) presenting evidence from production data.

The differences in accent type are reflected in varying degrees of perceived prominence, with falling/low accents being less prominent than rising/high accents (see Ladd and Morton, 1997). At the same time, the *position* of an accent has an impact on its strength or prominence, with the nucleus playing a special role. The nucleus is defined (both in the tradition of the British School, e.g. Crystal (1969), and in more recent autosegmental-metrical models of intonation, e.g. Ladd, 2008) as the last pitch accent in an intonation unit. It constitutes the only obligatory element in the phrase and is considered the *structurally* (phonologically) most important and thus most prominent element, whose location determines the interpretation of the phrase's information structure. However, the nucleus does not have to be the *phonetically* most prominent element (see e.g. Ayers, 1996). The metrical tree in Fig. 2 illustrates the proposed structural strength relations in a prosodic phrase: the nucleus takes the strongest position, since it is dominated by *s(tronger)*-nodes only, as opposed to the pre- and postnuclear elements, which represent terminal *w(eaker)*-nodes.

Experience with the annotation of spoken language shows that the detection and classification of nuclear accents is generally uncontroversial, whereas prenuclear accents are much more difficult to perceive and, in turn, to interpret. This is even more the case for postnuclear prominences, which – by definition – cannot be fully-fledged pitch accents but rather stresses or ''phrase accents'', i.e. edge tones with a secondary association to a stressed syllable, if present (see Grice et al., 2000, presenting evidence from various languages, including German). This hybrid status of phrase accents also

[5] An N400 is an electric potential in the brain measured in an electroencephalogram (EEG). It describes a centro-parietal negativity peaking around 400 ms after the onset of the target entity. A P600 describes a positive deflection peaking around 600 ms after target onset.

[6] The perception study was based on production data used in Röhr and Baumann (2010) and described in more detail in Section 5.1.

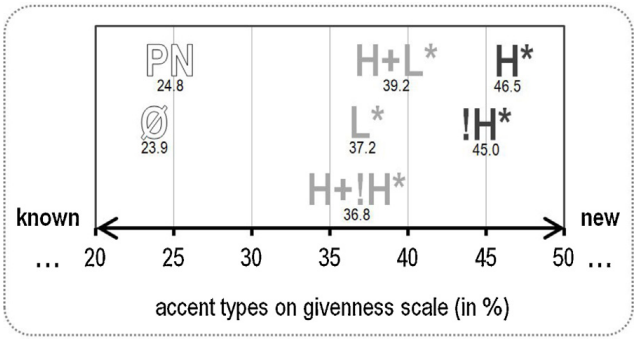

Fig. 1. Distribution of accent types (and positions: "PN" denoting prenuclear accents and "Ø" denoting deaccentuation) on a givenness scale in a perception experiment on German (Röhr and Baumann, 2011:1708).

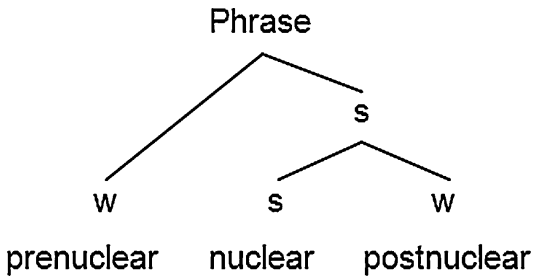

Fig. 2. Prominence structure of a prosodic phrase, illustrated by a metrical tree (following Liberman, 1975).

implies a lower level of prosodic prominence, i.e. postnuclear prominences can generally be regarded as weaker than prenuclear prominences, which, in turn, are weaker than nuclear prominences (cf. the classification in Baumann, 2012).

A recent perception experiment on German (Jagdfeld and Baumann, 2011) confirmed that listeners are much more sensitive to categorical prominence judgments in nuclear than in prenuclear (and postnuclear) position, which was also indicated by longer reaction times for prenuclear accents. These results are in line with findings by Baumann et al. (2007) as well as Féry and Kügler (2008), who showed that prenuclear accents are often placed for purely rhythmic reasons. That is, even fully given – non-contrastive – items may well be accented in prenuclear position (but not in nuclear position!), since the information structure of a sentence is not the only factor that determines its prosodic realisation.

For the data analysis in the present paper (Section 5), we will differentiate between three accent positions, i.e. nuclear accent, prenuclear accent and phrase accent, which can be assumed to decrease in their degree of prosodic prominence (given the evidence discussed), as well as lack of accentuation.[7]

3. The RefLex scheme

In this section, our goal is to investigate the different definitions which have been proposed for notions like *given*, *accessible* and *new*. In particular, we want to find out whether these definitions allow us to classify term expressions in texts in a satisfactory manner, in the sense that the classification should be intuitive for the annotator and, at the same time, avoid grouping together expressions that possess obvious pragmatic differences. In the present section, no reference will be made to prosody, in order to avoid any circular argumentation. Empirical findings about correlations between prosodic and meaning categories are the declared goal of the current investigation but they would become void if the information status categories themselves were tailored along the lines of prosodic features. Pragmatic classes should only make reference to meaning properties and, perhaps, to morphosyntactic features, like (in)definiteness, whose meaning contribution has been thoroughly investigated.[8] A second criterion for an annotation scheme, however, is that these meaning categories are defined with sufficient perspicuity, in order to enable even non-expert annotators to create concordant annotations. It is exactly because the RefLex scheme fulfils these two criteria that it provides a useful tool for research into the prosody of spoken data.

[7] In the analysis, we include *no accent* as a category of both variables *accent position* and *accent type*.

[8] Note that we do not annotate (in)definiteness separately but use it as a first criterion for defining the categories at the referential level. To annotators this has proven helpful, and it is in line with insights from the semantic-pragmatic literature.

Since the seminal work by Chafe (1976) and Prince (1981), it is commonly accepted that for the analysis of a text in informational terms it is not sufficient to distinguish only *given* and *new* information but to take at least a third intermediate class into account, sometimes described as *accessible* or *inferable* information. In the meantime, these classifications have undergone substantial modifications, for instance in work by Prince (1992), Lambrecht (1994), Eckert and Strube (2000), Nissim et al. (2004), Götze et al. (2007) or Riester et al. (2010). While the terminology in these accounts varies drastically, they at least agree that several subclasses must be taken into account. We refer to Baumann and Riester (2012) where a terminological and conceptual comparison of existing approaches is presented.

In the following we will briefly sketch the main issues for a satisfactory distinction of referential expressions and propose a two-dimensional classification system, the *RefLex* scheme.

3.1. Given versus New

A term is typically called given (active, old, textually evoked) if it has been mentioned earlier in the text. However, what has just been said can be interpreted in two different ways: either the *entity* referred to (represented by a discourse referent) has occurred before or the *lexical string* – or a part of it – is a repetition (it must be ensured, though, that the two strings are related in meaning, which excludes mere homonyms). While in (8) two different expressions can be said to corefer, in (9) two identical expressions refer to two separate entities.

(8) Smith was very optimistic. The polls showed a solid majority for the politician.

(9) After the holidays, John arrived in a new car, and also Harry had bought a new car.

As discussed in Section 2, both coreference and repetition or entailment of a concept seem to give rise to deaccentuation. While coreference is usually assumed to be a defining characteristic for givenness, entailment only occasionally is (for instance, in Schwarzschild, 1999). Of course, the two will frequently overlap, as they do in (10).

(10) A car was waiting in front of the hotel. I could see a woman in the car.

Another example involving repetition is the one in (11). This case is different from (9) since it involves two instances of the same generic indefinite *a cat*. This is another case of van Deemter's (1994) concept-givenness; compare example (6).

(11) A cat makes for a popular pet. Moreover, a cat is quite independent.

Generic expressions will not play a role in the Hypotheses formulated in the main sections of this article. However, they form an important aspect of the RefLex scheme without which some common distinctions found in natural data could not be made. We shall assume that the repetition of a generic item – definite or indefinite – does not count as coreference. This is what distinguishes generic expressions from specific ones which can establish coreference chains. Compare examples (11) and (8), and their analyses given below.

Furthermore, as mentioned in Section 2 in the discussion on concept-givenness, an anaphoric holonym or hypernym can count as given; as in (12), taken from van Deemter (1999:7). A related explanation is provided by Schwarzschild (1999) who claims that the previous mention of a hyponym *entails* a subsequent hypernym,[9] which is therefore given.

(12) Bach wrote many pieces for viola. He must have loved string instruments.

In order to account for these observations – while keeping the complexity of the annotation system at a low level – we propose to assign annotation labels at a referential as well as a lexical level. In a first step, we distinguish at the referential level whether some entity is specific and has occurred before **(r-given)** or is a specific entity introduced for the first time **(r-new)** or whether it is a generic concept **(r-generic)**. Secondly, we consider the lexical item itself and check whether it is a repetition, a holonym or hypernym **(l-given)** or a lexically unrelated word **(l-new)**. The examples mentioned above can thus all be analysed distinctively as follows[10]:

[9] More precisely, there is an entailment relation between the existentially type-shifted variants of the two expressions (Schwarzschild, 1999:155).

[10] In the following, r-labels are aligned with DPs, l-labels with nouns (or entire NPs): [the [politician]$_{\text{l-new}}$]$_{\text{r-given}}$. This mirrors the fact that the domain of the former is the entire referential expression (since only these introduce or take up a discourse referent), while the domain of l-labels are set-denoting expressions, cf. Baumann and Riester (2012).

(8′) Smith was very optimistic. The polls showed a solid majority for the politician.
r-given
l-new

(9′) After the holidays, John arrived in a new car,
r-new
l-new
and also Harry had bought a new car.
r-new
l-given

(10′) A car was waiting in front of the hotel. I could see a woman in the car.
r-new **r-given**
l-new **l-given**

(11′) A cat makes for a popular pet. Moreover, a cat is quite independent.
r-generic **r-generic**
l-new **l-given**

(12′) Bach wrote many pieces for viola. He must have loved string instruments.
r-generic **r-generic**
l-new **l-given**

3.2. Accessible

We claimed earlier that even when an entity cannot be analysed as *given* in one of the senses described above, there are still minor degrees to which it can count as activated. The most intuitive form of ‘‘semi-activation’’, ‘‘accessibility’’ (Chafe, 1976; Lambrecht, 1994) or ‘‘inferability’’ (Prince, 1981) applies if some item is a part of a previously mentioned object, i.e. a meronym, compare (13).

(13) I tried to open the door but the lock was rusty.

However, there are many cases which are intuitively similar to the one in (13) but for which the assumption of *inferability* and, therefore, *accessibility*, is unjustified. For instance, it seems simply wrong to assume that *the harpoon* in (14) is a weapon that anybody reasonably might have guessed or ‘‘inferred’’ upon simply hearing about the murder incident. Neither is it a meronym of some expression in the context.

(14) John was murdered recently. The harpoon lay nearby.

Rather, what the definite expressions in (13) and (14) have in common is that they are context-dependent expressions (anaphors) which, however, do not possess a coreferential antecedent. Such cases have been called *bridging anaphors* or *associative anaphors* (Löbner, 1998; Vieira and Poesio, 2000). They adhere to the so-called *given-new contract* (Clark and Haviland, 1977) in the sense that the speaker marks an expression as definite and therefore identifiable. The hearer, unable to identify the referent itself in the discourse context, subsequently constructs a ‘‘bridge’’ to link the anaphor to previously mentioned material. At the referential level, we therefore add the label **r-bridging** to our inventory while *accessibility* is handled at the lexical level: **l-accessible** describes a lexical relation of subordination (hyponymy or meronymy).[11] Accordingly, sentence (13) will be analysed as follows:

(13′) I tried to open the door but the lock was rusty.
r-bridging
l-accessible

The opposite order of items will be annotated as in (13b).

(13b) I put the key in the lock but the door wouldn’t open.
r-bridging
l-given

[11] Remember that the opposite sequence, e.g. a holonym preceded by one of its meronyms was classified by means of the label l-given.

The analysis for (14) is (14′).

(14′) John was murdered recently. The harpoon lay nearby.
r-bridging
l-new

We furthermore assume that obvious semantic relations between two expressions of which the latter is indefinite can also be captured using our previously established label inventory, compare (15).

(15) When I came into my office, a window stood open.
r-new
l-accessible

3.3. Discourse-New versus Hearer-New

In Prince (1981), (discourse-)new items are distinguished into *brand-new* and *unused* ones, the former are *hearer-new*, the latter are *known* or *familiar*. Similar classifications exist in Lambrecht (1994:109), in which known entities are called *inactive* (and therefore rank below *accessible* ones on a cognitive scale), while *unknown* entities are treated outside the activation scale and are called *unidentifiable*. Another well-known approach is the *givenness hierarchy* by Gundel et al. (1993), in which, however, terminology is applied differently. *Known* entities are called *familiar*[12]; *unknown* ones are classified as *identifiable, not familiar*. In more recent approaches, this has led to the decision to only count *hearer-new* items as *new*, while *known* ones are allocated to the intermediary class of *accessible* (Götze et al., 2007) or *mediated* (Nissim et al., 2004) items. Still, the common denominator of all these approaches is that they distinguish between two types of non-anaphoric definite expressions, *known* ones like in (16) and *unknown* ones like in (17).

(16) President Barack Obama delivered a brilliant speech in Tucson.

(17) The person who stole my wallet yesterday was wearing glasses.

There are two problems here. The first one is that postulating one big class of *brand-new (unknown, unfamiliar, unidentifiable)* items will group cases like in (17) together with indefinites like in (18).

(18) A man came into Rick's Café.

There may be a cross-linguistic reason for assorting the expressions in (17) and (18). Many languages do not distinguish between a definite and an indefinite article. Yet, for English or German it is unclear how such a move should be justified from a semantic perspective since it is commonly assumed that definite articles trigger presuppositions of uniqueness within some domain (Heim, 1982; van der Sandt, 1992; Beaver and Geurts, 2011) while indefinite articles do not.

A more severe problem is the reliability of such annotations from a practical perspective. Quite often, non-anaphoric definites, like proper names or complex definite descriptions, cannot easily be classified as *known* or *unknown* like this was possible for the clear cases (16) and (17), since the crucial criterion for such a classification is the knowledge of the intended audience (cf. Riester, 2008). The latter is often a factor of uncertainty, which may lead to inhomogeneous annotations. For instance, the underlined person in (19) is known to a local group of people but unknown to a larger audience.

(19) Cologne Archbishop Meisner has denied homosexual couples the right to adopt children.

We think that this widespread uncertainty should not be taken as a criterion for the decision of whether to classify an expression as *new* or *accessible*. We also think that definites like in (17) and indefinites like in (18) should be kept apart in the annotation process. We therefore postulate the following rules: **r-new** only applies to (non-generic) indefinites, while discourse-new, non-anaphoric definites should be grouped together under the label **r-unused**,[13] as for (16) (which receives the same label combination like (17) and (19)).

[12] More precisely: familiar but not activated.

[13] This conflation deviates from Prince's (1981) use of the notion *unused*, which only applies to hearer-known expressions. By contrast, in Baumann and Riester (2012), we distinguish between *r-unused-known* and *r-unused-unknown*, if possible.

(16′) President Barack Obama delivered a brilliant speech in Tucson.
r-unused
l-new

But:

(18′) A man came into Rick's Café.
r-new
l-new

3.4. The RefLex scheme (simplified): summary and overview

In Tables 1 and 2, an overview of the scheme – divided into a referential and a lexical level – is presented. The overview is a simplification of a more comprehensive account and describes only those labels which play a role in the empirical study below. For the entire scheme, consult Baumann and Riester (2012).

Table 1
Labels for the annotation of discourse referents; ''r-'' indicates the referential level.

Units: Definite DPs	
r-given	anaphor corefers with antecedent in previous discourse
r-bridging	anaphor can be resolved to non-coreferring antecedent or within a described scenario
r-unused	discourse-new, non-anaphoric definite expression referring to an item which is generally known or identifiable from its own linguistic description
Units: Definite or Indefinite DPs	
r-generic	abstract or generic item
Units: Indefinite DPs	
r-new	specific or existential indefinite introducing a new referent

Table 2
Labels for the annotation of nouns denoting discourse referents; ''l-'' indicates the lexical level. Pronouns do not receive l-labels.

Units: Nouns (also NPs and other set-denoting expressions)	
l-given	noun is a repetition (excluding homonyms) or a synonym, or it is lexically superordinate to a previous noun (markable is a hypernym, holonym, superset)
l-accessible	noun is lexically subordinate to previous noun (markable is a hyponym, meronym, subset)
l-new	noun not related to another noun within last five clauses or intonation phrases[a]

[a] The decision to choose five clauses is to some extent arbitrary but it serves as a clear rule for transcribers in the annotation process. In fact, it is a matter of debate in the literature how far away a referent's last mention may be to still count as given, and to what extent recency effects can be thought of as a linear function of distance. According to Chafe (1987), a referent's deactivation from an earlier state may have at least two reasons: Apart from a simply gradual (or linear) decay mechanism, a referent may become less accessible due to interference from competing referents mentioned in the intervening discourse (which is non-linear). The aspect of non-linearity is emphasised by Clark and Sengul (1979) who found a significantly higher availability of a referent mentioned in the previous clause compared to a referent mentioned two clauses back, while there was no significant effect between referents from two clauses back and referents from three clauses back.

It is important to note that the RefLex scheme does not (and is not meant to) cover all relevant aspects which may influence an item's prosodic realisation. That is, RefLex is just a part of a more comprehensive annotation scheme which will at least include labels for configurations which *elicit alternatives*, such as focus-sensitive particles or overtly contrastive expressions (Rooth, 1992; Selkirk, 2008; Beaver and Velleman, 2011; Riester and Baumann, 2013). We will not discuss these features in detail here, but we are well aware that focus/contrast often overrides the prosody that would be assigned to an item due to its information status alone (see e.g. Baumann (2006) and the comments in the conclusions of the present article). Furthermore, the present investigation does neither take syntactic positions nor word order phenomena into account, which may also be relevant for an utterance's prosodic surface form.

4. Hypotheses

For the rest of the paper, we want to investigate *empirically* which combinations of information status levels serve as triggers for which intonational categories. The hypotheses we want to test in two types of data, namely read speech

(Section 5.1) and spontaneous monologues (Section 5.2) are motivated by the discussion of the examples given in Section 2 as well as by the taxonomy presented in Section 3. The numbering of the respective examples is repeated for each hypothesis. We regard this sample of combinations of referential and lexical givenness categories found in the literature as particularly clear cases which are revealing with respect to the role prosody can play in information packaging. However, we do not simply adopt the claims of accentuation versus deaccentuation made in the literature. Instead, we assume *relative* differences in the intonational marking of information status, indicated by different *distributions* and/or *probabilities* of prosodic categories. The categories we are looking at are *accent position* (nuclear accent, prenuclear accent, phrase accent[14] or no accent – assuming a decrease in prominence from left to right; see Section 2) and *accent type* for nuclear and prenuclear pitch accents (H*, !H*, Early Peak accents, or L* – assuming a decrease in prominence from left to right; see Section 2) as opposed to lack of pitch accent (comprising phrase accents and complete deaccentuation).

We generally hypothesise that the ''newer'' a referring expression is (both at the referential and the lexical level) the more it is made prosodically prominent by a speaker. The following sub-hypotheses propose differences in the prominence relation between specific (and supposedly most revealing) RefLex combinations. We start with a comparison of the extreme cases, i.e. combinations of new labels at both levels with given labels at both levels.

Hypothesis I. New referents (r-new) encoded by new lexical items (l-new) (cf. *car* example (2) in Section 2) are marked as more prominent than given referents (r-given) encoded by given lexical items (l-given) (cf. *dog* example (3).

The next two hypotheses are expected to be weaker than Hypothesis I. In each of them, we keep one of the levels constant, selecting only *new* expressions, while comparing *givenness* with *newness* at the other level.

Hypothesis II. New referents (r-new) encoded by new lexical items (l-new) (cf. *car* example (2)) are marked as more prominent than new referents (r-new) encoded by given lexical items (l-given) (cf. *dog* example (5)).

Hypothesis III. New referents (r-new) encoded by new lexical items (l-new) (cf. *car* example (2)) are marked as more prominent than given referents (r-given) encoded by new lexical items (l-new) (cf. *butcher* example (4)).

The following hypothesis pertains to the distinction *r-new* vs. *r-unused* which is specific to the RefLex scheme. We investigate whether, from a prosodic perspective, it makes sense to distinguish between indefinite (r-new) and definite (r-unused) discourse-new expressions.

Hypothesis IV. New referents (r-new) encoded by new lexical items (l-new) (cf. *car* example (2)) are marked as more prominent than unused referents (r-unused) encoded by new lexical items (l-new) (cf. *president Obama* example (16) in Section 3).

The next two hypotheses mirror Hypotheses II and III. In V and VI we keep *givenness* constant at either the referential or the lexical level and vary the complementary level, respectively. As before, we expect modifications at each level to yield changes in prosodic prominence, which deviates from the claims made in the literature suggesting that givenness at either level generally leads to deaccentuation (cf. examples (4) and (5) above).

Hypothesis V. Given referents (r-given) encoded by new lexical items (l-new) (cf. *butcher* example (4)) are marked as more prominent than given referents (r-given) encoded by given lexical items (l-given) (cf. *dog* example (3)).

Hypothesis VI. New referents (r-new) encoded by given lexical items (l-given) (cf. *dog* example (5)) are marked as more prominent than given referents (r-given) encoded by given lexical items (l-given) (cf. *dog* example (3)).

The final hypothesis deals with accessible or inferable information realised on bridging anaphors. We investigate whether there is an increase in prosodic prominence between superordinate and subordinate bridging anaphors.

Hypothesis VII. Anaphoric but non-coreferential expressions (r-bridging) encoded by items which are lexically subordinate to an antecedent (l-accessible) (cf. *drawers* example (7)) are marked as more prominent than anaphoric but non-coreferential expressions (r-bridging) encoded by items which are lexically superordinate to an antecedent (l-given) (cf. *desk* example (6)).

[14] Phrase accents are defined here as postnuclear prominences, not as boundary tones.

5. Data analysis

5.1. Read speech

5.1.1. Method

For the present study, we reanalysed the speech material elicited in Röhr and Baumann (2010) (mentioned in Section 2) using the RefLex scheme. The test material consisted of ten different stories (divided in three separate paragraphs) in which ten different target words occurred in four different types of information status each. An example paragraph including the annotations on the referential and lexical levels is given in Table 3. Referring expressions are underlined. Labels for the lexical level are aligned with the respective nouns.

Table 3
Example reading material for the target word *Banane* ('banana'), annotated according to the simplified RefLex scheme used in the present paper.

„Was	**hätten**	**Sie**	**gerne?".**	**„Ich**	**nehme**	**die**	**Banane**	**mit.",**	**antwortet**
		r-given		r-given		r-unused			
		ϕ		*ϕ*			l-new		
Thomas	**dem**	**Obsthändler.**	**Normalerweise**	**ernährt**	**er**	**sich**	**sehr**	**ungesund**	**und**
r-unused	r-bridging				r-given	r-given			
l-new		l-new			*ϕ*	*ϕ*			
isst	**zwischendurch**	**ständig**	**Süßigkeiten.**	**Ausserdem**	**treibt**	**er**	**fast**	**nie**	**Sport**
			r-generic			r-given			r-generic
			l-new			*ϕ*			l-new
und	**wenn**	**doch,**	**dann**	**am**	**liebsten**	**Minigolf.**	**Er**	**steckt**	**sich**
						r-generic	r-given		r-given
						l-accessible	*ϕ*		*ϕ*
die	**Banane**	**ein.**	**Lecker**	**sieht**	**die**	**Banane**	**aus.**	**Vielleicht**	**wird**
r-given					r-given				
	l-new					l-given			
er	**demnächst**	**öfter**	**welche**	**kaufen.**					
r-given			r-generic						
ϕ			*ϕ*						

"*What would you like?*" "*I'll take the banana (along)*", *Thomas answers the greengrocer. He usually lives on unhealthy food and always eats sweets between meals. He hardly ever practices sport, and if he does he prefers mini golf. He pockets the banana. The banana looks delicious. Maybe he'll buy them more often in future*.

The texts were read by ten native speakers of German (seven female, three male), aged between 22 and 31. All of them originated from the area around Cologne and Düsseldorf (North Rhine-Westphalia). Before the acoustic recordings, which were done in a quiet room, each subject was asked to read through the material thoroughly in order to guarantee full comprehension. After that, the subjects' task was to read out the texts in a contextually appropriate manner. The three contexts were presented on separate file cards in pseudo-randomised order. They were repeated three times by each subject, adding up to 450 sentences per speaker. The semantic and prosodic labelling was done with Praat (Boersma and Weenink, 1996) by three independent annotators.

In our reanalysis, not only the original target words are annotated but all referring expressions in the ten stories. So far, we annotated the complete data set of two speakers plus a selection of 23 sentences produced by ten different speakers. The data consist of 638 intonation phrases and 4673 words, including 1707 referents. The results reported in the following section are based on a selection of 807 referential expressions whose classification is (at least for the most part) relevant for testing our hypotheses. That is, pronouns (which are not marked at the lexical level) as well as less frequent and/or less relevant label combinations were excluded.

5.1.2. Results

Fig. 3 shows the relative distribution of accent *positions* for the most relevant RefLex combinations. The combinations are grouped according to their lexical label, and within each group arranged from new to given at the referential level.

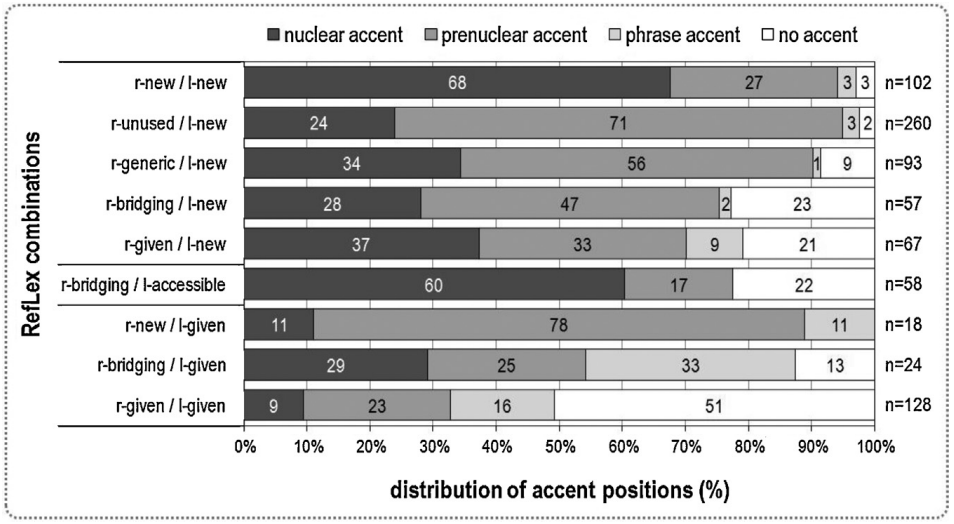

Fig. 3. Distribution of accent positions for selected RefLex combinations (in percent) and absolute numbers of occurrences in the read speech corpus (807 tokens in total).

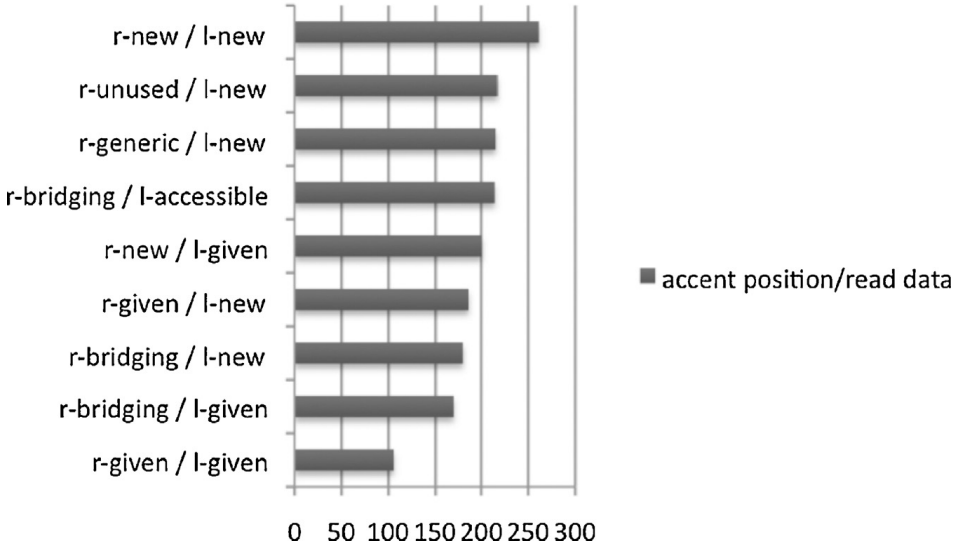

Fig. 3a. Prominence scores for accent position (P_{AP}) in the read corpus.

Before discussing the results of Fig. 3 against our hypotheses, we first define a measure for *prosodic prominence* which enables us to compare the relative prominence of label combinations in an intuitive way. The measure certainly is *ad hoc* but also transparent since it is derived from the evidence discussed in the literature and at the end of Section 2 suggesting differences in the degree of nuclear, prenuclear and postnuclear prominences. Fig. 3a shows the label combinations from Fig. 3 ranked according to the following prominence score P_{AP} for accent position:

$$P_{AP} = \text{ratio of nuclear accents} \cdot 3 + \text{ratio of prenuclear accents} \cdot 2 + \text{ratio of phrase accents}$$

(Deaccented expressions receive a weight of zero. A maximal score of 300 would indicate that all instances of a label combination are marked by means of a nuclear accent.)

The overall results presented in Fig. 3a confirm our general hypothesis saying the ''newer'' a referring expression the higher its degree of prominence, indicated here by the *position* of an accent (including deaccentuation). The only obvious exception is the rather low prominence of the combination r-bridging/l-new indicated by a large number of deaccentuations (23%, see Fig. 3), which came as a surprise.

Let us turn to the specific hypotheses. Hypothesis I is clearly confirmed, since fully new items (i.e. items which are new at both levels) are marked by pitch accents in 95% of the cases (68% nuclear accents), whereas fully given items are only pitch accented in 32% (9% nuclear accents), and deaccented in 51% of all cases. A chi square test revealed a highly significant difference between the distributions of accent positions ($\chi^2(3, n = 230) = 108.75$, $p < 0.001$).

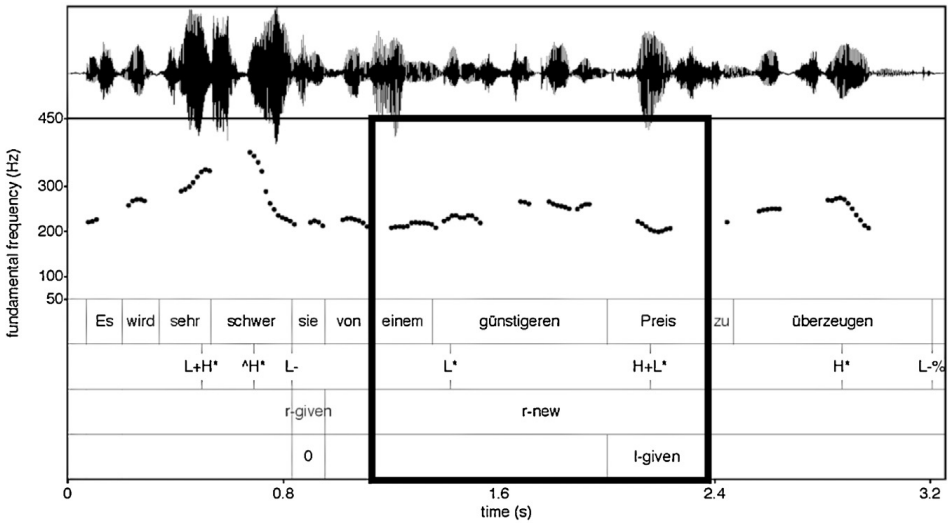

Fig. 4. Pitch contour and label tiers of the read utterance *Es wird sehr schwer, sie von einem günstigeren Preis zu überzeugen* ('It will be very difficult to convince her of a cheaper price').

Hypotheses II and III are confirmed as well, showing that the combination r-new/l-new is made more prominent by the speakers than combinations of new and given labels, indicated by more nuclear accents and less deaccentuations (both comparisons are significant: r-new/l-new vs. r-new/l-given $\chi^2(3, n = 120) = 23.03$, $p < 0.001$; r-new/l-new vs. r-given/l-new $\chi^2(3, n = 169) = 22.96$, $p < 0.001$). Fig. 4 shows a typical example of a referentially new but lexically given item in our corpus of read speech. Labelling tiers include the words spoken, tone labels for pitch accents, break indices and boundary tones (following GToBI), labels indicating information status at the referential and lexical levels as well as the pitch (F0) contour of the utterance. The word in question, *Preis* ('price'), was mentioned before but denoted a different referent. It is marked by a prenuclear accent which is by far the most common accent position found for the r-new/l-given combination (in this case, an Early Peak accent).

Interestingly, we found hardly any difference in their mean prosodic prominence when comparing the two reverse combinations of givenness and newness (see Fig. 3a). However, while r-given/l-new triggers more nuclear accents but also more cases of deaccentuation than r-new/l-given, the latter is predominantly marked by prenuclear accents. Thus, the converse effects of cognitive activation at both levels seem to counterbalance each other to a large extent.[15]

We did find a significant difference between indefinite (r-new) and definite (r-unused) discourse-new expressions in terms of accent position, which confirms Hypothesis IV ($\chi^2(3, n = 362) = 63.95$, $p < 0.001$). This difference is primarily indicated by a larger percentage of nuclear accents of r-new/l-new items in contrast to more prenuclear accents (which are structurally less prominent and potentially determined only by the speech rhythm; see end of Section 2) on r-unused/l-new expressions. In other words, the prosodic results justify our decision to keep discourse-new definites and indefinites apart.

Hypotheses V and VI are confirmed as well, revealing that only *two-level* givenness predominantly leads to deaccentuation, while newness at the lexical or the referential level mostly triggers pitch accents (r-given/l-new: mostly nuclear and prenuclear accents; r-new/l-given: mostly prenuclear accents). Both comparisons were highly significant (r-given/l-new vs. r-given/l-given $\chi^2(3, n = 195) = 31.01$, $p < 0.001$; r-new/l-given vs. r-given/l-given $\chi^2(3, n = 146) = 24.93$, $p < 0.001$).

Finally, we found a clear difference between subordinate and superordinate bridging anaphors, the latter being less prominent ($\chi^2(3, n = 82) = 23.93$, $p < 0.001$). Thus, Hypothesis VII is confirmed. While 60% of all r-bridging/l-accessible expressions are marked by nuclear accents, a relatively large number of items labelled r-bridging/l-given were marked by a phrase accent (33%), an example of which is shown in Fig. 5 (phrase accents are symbolised by a star in brackets in GToBI). Here, the referring expression *Schule* ('school') can be bridged from *Lehrer* ('teacher'), and is a holonym of this antecedent.[16]

[15] Note that we refrained from adding another hypothesis for the distinction between r-new/l-given and r-given/l-new, since it is statistically unsound to postulate the *lack* of a difference between two variables (i.e. the null hypothesis).

[16] We believe that, in cases of bridging, the question which word is the actual *antecedent* plays a minor role, since we view bridging anaphors as expressions whose interpretation depends on some scenario context rather than on a single predecessor. In many cases, this scenario may be evoked by a single antecedent expression but this need not be the case in general. At the lexical level, however, the existence and choice of an antecedent word plays a crucial role for determining the label. The fact that these two facts can be kept apart is central to the RefLex scheme.

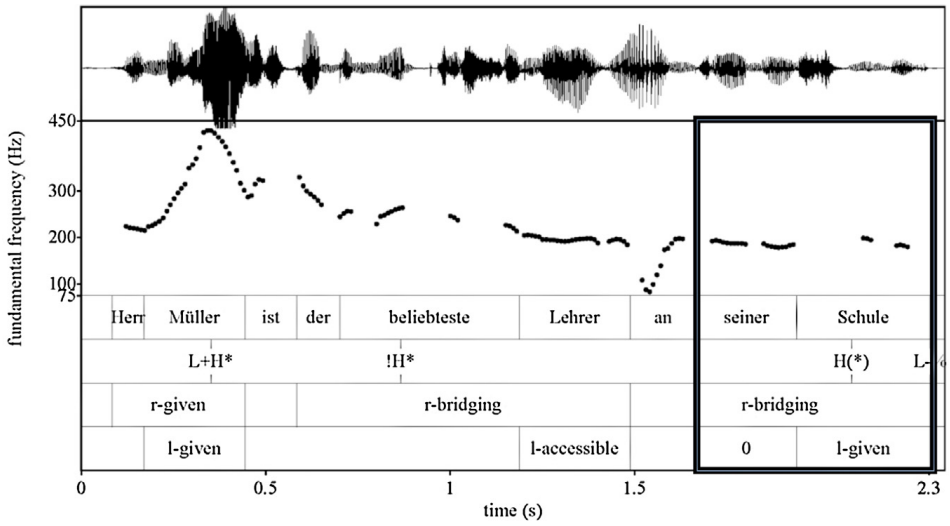

Fig. 5. Pitch contour and label tiers of the read utterance *Herr Müller ist der beliebteste Lehrer an seiner Schule* ('Mr. Müller is the most popular teacher at his school').

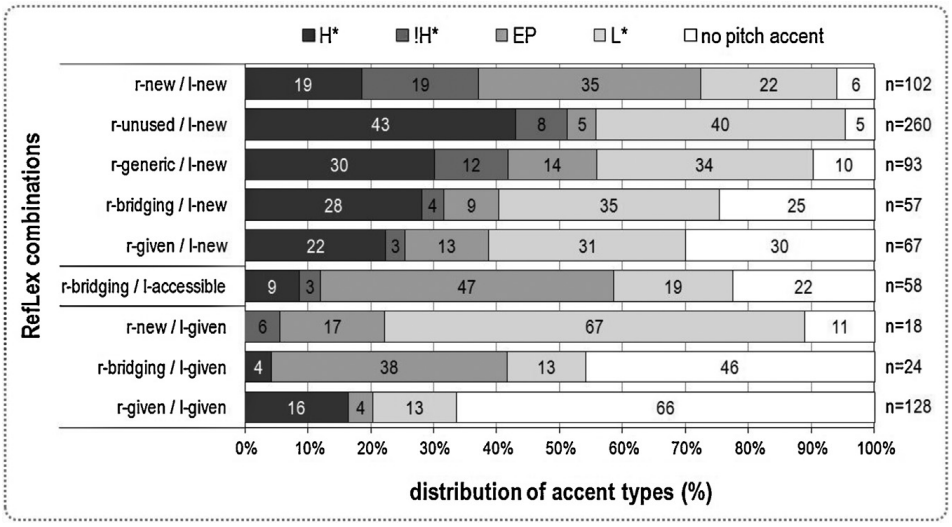

Fig. 6. Distribution of accent types for selected RefLex combinations (in percent) and absolute numbers of occurrences in the read speech corpus (807 tokens in total).

The results for the distribution of accent positions do not warrant the assumption that l-accessible information should be marked less prominently than l-new information. This, however, is confirmed after all as we look at the distribution of accent *types*.

The most frequent general accent type in the read data was H* (39%), followed by L* (35%) and Early Peaks (16%).[17] More specifically, L + H* was the most common prenuclear accent type (33%), and H + L* (an Early Peak accent) the most frequent nuclear accent type (29%). The combination of both types of accent proved to be the predominant reading intonation in our data. Fig. 6 gives the relative distribution of accent types for the same RefLex combinations as above. In this figure, we include the category ''no pitch accent'' (comprising phrase accents and cases of deaccentuation)[18] as an ''accent type'' for reasons of clarity. By doing so, we make sure to have the same numbers of tokens as in Fig. 3 and do not miss the information on the general ratio of accentuation for each label combination.

[17] H* and L* are used as classes here, comprising several single accent types with the same starred element, e.g. L + H* and ^H* count as H*.

[18] A note on the status of phrase accents: Recall that a phrase accent is a hybrid of an accent and a boundary tone, and is thus secondary both in terms of phonological status and prosodic prominence. It is more readily definable in terms of accent *position* (since it always occurs postnuclearly – although *prominence position* would be a more appropriate term) than of accent *type*, given that a phrase accent simply is no fully-fledged pitch accent, primarily due to its lack of a considerable tonal movement on the associated stressed syllable (there may still be a tonal target on this syllable, though; see Grice et al., 2000). This is why we include phrase accents as a category of prominence in Figure 3 but not as a pitch accent category in Figure 6.

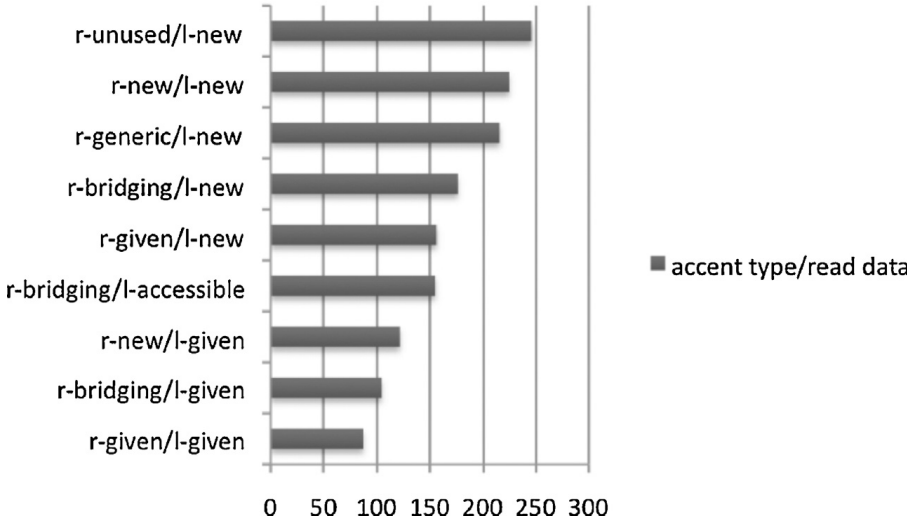

Fig. 6a. Prominence scores for accent type (P_{AT}) in the read corpus.

We define a second prominence measure P_{AT}, which integrates the findings from Section 2 concerning the assumed communicative impact of different accent types. Fig. 6a shows the ranking according to the prominence measure for accent types, which is defined as:

P_{AT} = ratio of H* · 4 + ratio of !H* · 3 + ratio of EP · 2 + ratio of L*

(Lack of pitch accent amounts to a weight of zero. A maximal ratio of 400 would indicate that all instances of a label combination are accented by means of H*.)

Fig. 6a reveals an orderly sequence of prominence degrees corresponding to the information status at the *lexical* level, i.e. ranging from l-new through l-accessible to l-given. Within these three groups, prominences are ordered according to the information status at the *referential* level. The only unexpected finding is that the combination r-unused/l-new is marked as more prominent than r-new/l-new, which differs from the order of prominence on the basis of accent positions presented in Fig. 3a. However, the reason for the high prominence score of r-unused/l-new lies exactly in the interaction with accent position: the large number of H*/prominent accents can be explained by the high ratio of prenuclear accents found with this label combination (71%, cf. Fig. 3), for which (L +)H* was the most common accent type. Actually, unused referents often occur as topics in initial sentence position in the corpus, confirming the widespread assumption that a rising pitch accent is the usual topic accent in German, although often a L* + H accent (instead of L + H*) is considered to be the default topic marker (see e.g. Uhmann, 1991).

Another difference between Figs. 6a and 3a is the reversed location of r-bridging/l-new and r-bridging/l-accessible, the latter being less prominent with respect to accent type. In fact, lexical accessibility is marked by a particularly large number of Early Peak accents (47% of all cases; cf. Fig. 6).

Despite the desirable results shown in Fig. 6a, the distribution of accent types only partly matches our expectations. While in the group of lexically new items most combinations show a high percentage of high or rising accents, as hypothesised, the combination r-new/l-new, which is expected to be prosodically most prominent, is mostly marked by (less prominent) Early Peak accents (in 35% of the cases; cf. Fig. 6).[19] Moreover, all l-new combinations are characterised by high percentages of L* accents, which qualifies the claims made in previous studies (e.g. Kohler, 1991; Pierrehumbert and Hirschberg, 1992; see Section 2). According to the general hypothesis, the l-given combinations show fewer high accents and more Early Peak and low accents. An exception is r-given/l-given, which is surprisingly often marked by H* accents (in 16% of the cases; cf. Fig. 6).[20]

The somewhat unexpected distributions of accent types only lead to the rejection of Hypothesis IV, though: r-unused/l-new expressions were found to be marked as more prominent than r-new/l-new expressions ($\chi^2(4, n = 362) = 78.63$, $p < 0.001$, opposite direction as hypothesised). Hypotheses I–III, V and VI were confirmed.[21] No significant difference

[19] This result is influenced by the fact that two thirds of r-new/l-new items occur in nuclear position (cf. Figure 3), for which H + L* has been found to be the most common type of accent.

[20] Again, it has to be taken into account that two-thirds of the accents are prenuclear, for which L + H* was most common.

[21] Hypothesis I: $\chi^2(4, n = 230) = 110.23$, $p < 0.001$; Hypothesis II: $\chi^2(4, n = 120) = 18.16$, $p < 0.01$; Hypothesis III: $\chi^2(4, n = 169) = 32.12$, $p < 0.001$; Hypothesis V: $\chi^2(4, n = 195) = 28.51$, $p < 0.001$; Hypothesis VI: $\chi^2(4, n = 146) = 45.49$, $p < 0.001$.

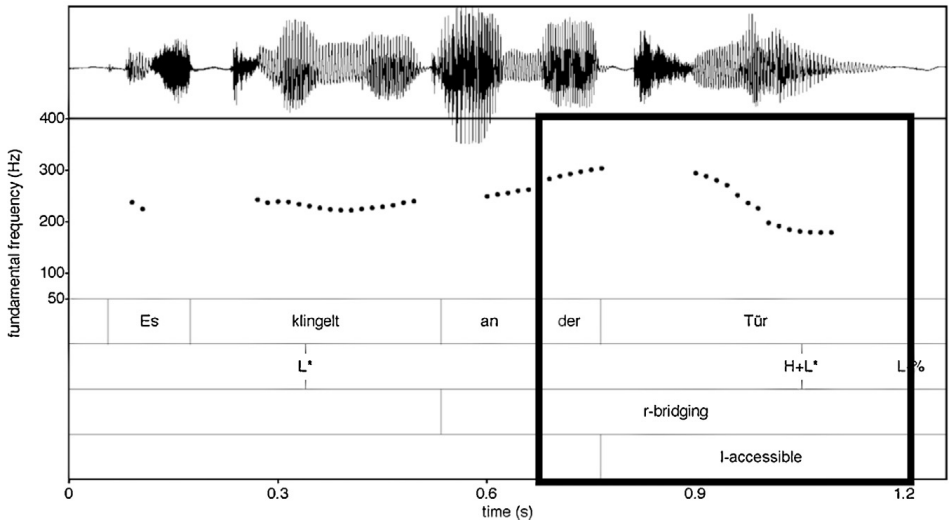

Fig. 7. Pitch contour and label tiers of the read utterance *Es klingelt an der Tür* ('The door bell rings').

Table 4
Hypotheses of prominence relations between RefLex combinations and results of chi square (χ^2) tests in the read corpus data; ">" indicates higher degree of prominence; the *p* value in brackets indicates a difference in the opposite direction than proposed by the hypothesis.

		Accent position	Accent type
Hypothesis I	r-new/l-new > r-given/l-given	$p < 0.001$	$p < 0.001$
Hypothesis II	r-new/l-new > r-new/l-given	$p < 0.001$	$p < 0.01$
Hypothesis III	r-new/l-new > r-given/l-new	$p < 0.001$	$p < 0.001$
Hypothesis IV	r-new/l-new > r-unused/l-new	$p < 0.001$	($p < 0.001$)
Hypothesis V	r-given/l-new > r-given/l-given	$p < 0.001$	$p < 0.001$
Hypothesis VI	r-new/l-given > r-given/l-given	$p < 0.001$	$p < 0.001$
Hypothesis VII	r-bridging/l-accessible > r-bridging/l-given	$p < 0.001$	n.s.

could be found between the two combinations with r-bridging (Hypothesis VII), both of which were for a major part (37% and 48%, respectively; cf. Fig. 6) marked by Early Peak accents. A typical example of a bridged and lexically subordinate item is given in Fig. 7. Here, the referent *Tür* ('door'), which is a meronym of the previously mentioned *Wohnung* ('apartment'), is prosodically marked by a H + L* pitch accent in nuclear position.

Table 4 provides a summary of our results for read speech.

5.2. Spontaneous monologues

5.2.1. Method

We analysed six spontaneous monologues, which are part of a larger corpus on German currently being compiled for an investigation of various types of spoken data. The monologues annotated were produced by six native speakers originating from North Rhine-Westphalia (like the speakers of the read data set) and aged between 27 and 42. Their only instruction was to tell a story of their choice for no longer than five minutes. The six stories, which were digitally recorded in a quiet room, consist of 374 intonation phrases comprising 3008 words and 819 referents (see also Baumann and Riester, 2010). The labelling at various linguistic levels was done by two independent annotators. For the present paper, we concentrate on the consensus annotation of the information status of nominal categories at the referential and lexical levels, and their phonological marking in terms of accent position and accent type. As with the read data presented above, the results reported in the following section are restricted to a selection (218 referents) of the expressions which are relevant for the hypotheses.

5.2.2. Results

Fig. 8 gives an overview of the distribution of accent positions for the combinations of RefLex categories that are dealt with in the present paper. Note that the three combinations r-bridging/l-accessible, r-new/l-given and r-bridging/l-given, which we discussed with respect to the read corpus, had to be left out due to the small number of occurrences (i.e. less than 10 cases).

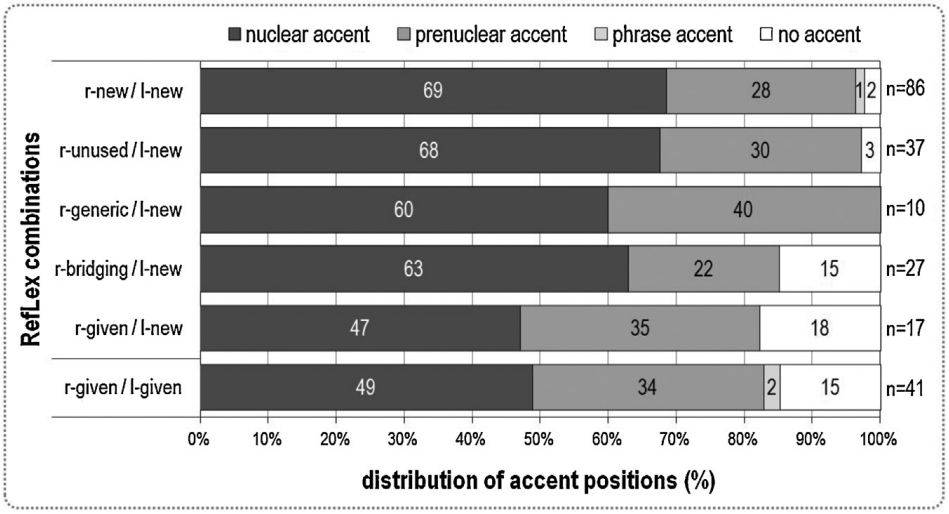

Fig. 8. Distribution of accent positions for selected RefLex combinations (in percent) and absolute numbers of occurrences in the spontaneous speech corpus (218 tokens in total).

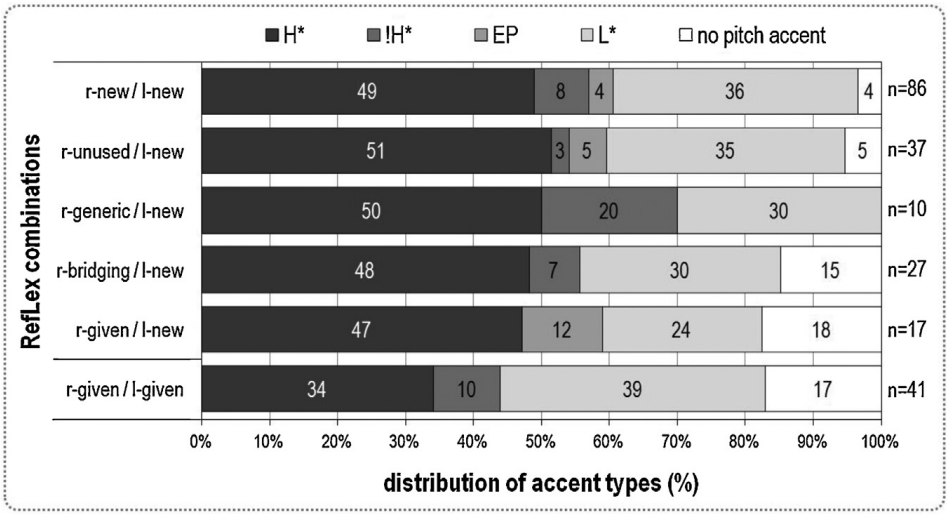

Fig. 9. Distribution of accent types for selected RefLex combinations (in percent) and absolute numbers of occurrences in the spontaneous speech corpus (218 tokens in total).

Generally, all combinations are for the largest part marked by nuclear accents. Nevertheless, a slight decrease in prominence within the group of combinations containing l-new (from r-new to r-given) can be observed, which is predominantly due to an increase in the number of deaccentuations. Interestingly, though, the distribution of accent positions for fully given items (r-given/l-given) is very similar to the distribution for r-given/l-new elements. Thus, our general hypothesis proposing increased prominence from given to new items at both levels can at most be claimed to be confirmed in tendency. We have refrained from depicting the prominence scores since they do not add much to the data in Fig. 8.

This rather weak result for the spontaneous data set also becomes obvious when we look at the specific hypotheses: only the comparison of fully new with fully given expressions yields a significant difference of accent positions ($\chi^2(3, n = 127) = 9.08$, $p < 0.05$), confirming Hypothesis I. The other three hypotheses that can be tested (only Hypotheses III–V can be based on a sufficient number of cases) have to be rejected, since no significant differences between the RefLex combinations were found.[22]

Turning to the analysis of accent *types* in the spontaneous data, we observe that H* is the most frequent class (53%) followed by L* (35%) (as a percentage of all fully-fledged pitch accents). A plain H* is the most common prenuclear accent

[22] By looking at the distributions, the lack of a significant difference in a chi square test between r-new/l-new and r-given/l-new (Hypothesis III) came as a surprise. The reason might be the relatively small number of r-given/l-new cases ($n = 17$).

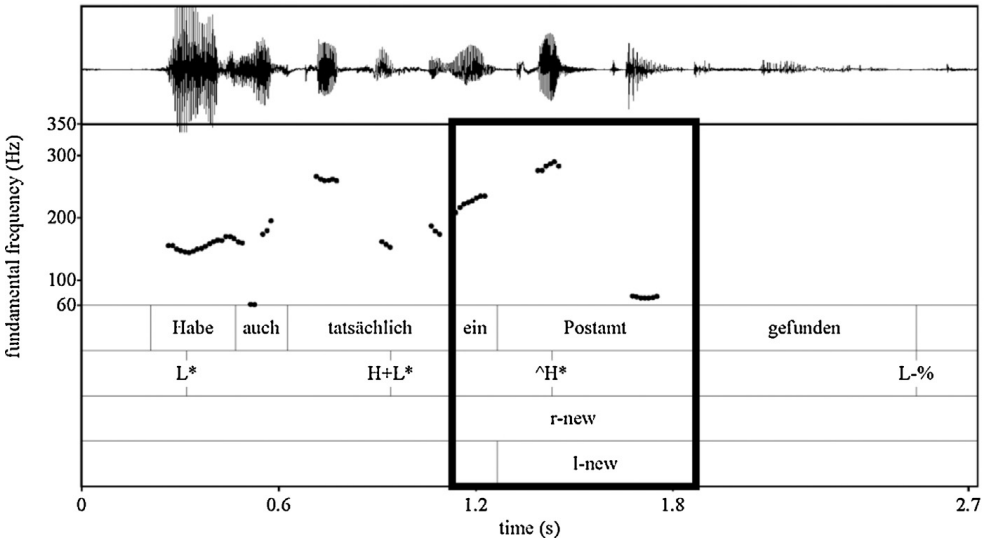

Fig. 10. Pitch contour and label tiers of the spontaneous utterance *Habe auch tatsächlich ein Postamt gefunden* ('and I actually found a post office'); the framed DP *ein Postamt* ('a post office'), labelled as r-new and l-new, is marked by an (upstepped) high accent.

Table 5
Hypotheses of prominence relations between RefLex combinations and results of chi square (χ^2) tests in the spontaneous corpus data; ">" indicates higher degree of prominence.

		Accent position	Accent type
Hypothesis I	r-new/l-new > r-given/l-given	$p < 0.05$	n.s.
Hypothesis II	r-new/l-new > r-new/l-given	–	–
Hypothesis III	r-new/l-new > r-given/l-new	n.s.	n.s.
Hypothesis IV	r-new/l-new > r-unused/l-new	n.s.	n.s.
Hypothesis V	r-given/l-new > r-given/l-given	n.s.	n.s.
Hypothesis VI	r-new/l-given > r-given/l-given	–	–
Hypothesis VII	r-bridging/l-accessible > r-bridging/l-given	–	–

(34%) and L* the most frequent nuclear accent type (42%). The nuclear L* accents are often followed by a high boundary tone, combining for a rising nuclear movement. Such a movement serves as a continuation rise, which is typical of spontaneous storytelling in German (as an example, see Fig. 12 in Section 6 below). Similar to the read paragraphs, the most common single accent type was L + H*, which occurred in 30% of all cases. In contrast to the read data, however, Early Peak accents were rare (5%). See Fig. 9 for the relative distribution of accent types assigned to the RefLex combinations.

The figure reveals that there is little variation in the overall accent type distribution, especially within the group of combinations with lexically new expressions. The only modification seems to be that generic and bridged referents are not marked by Early Peak accents and that plain downsteps do not occur with given referents. On the whole, the general hypothesis of a correlation between increased newness and increased prominence is not supported by these data.

Nevertheless, Fig. 10 gives an example in favour of our hypothesis, namely a referentially and lexically new item (*Postamt*, 'post office') which is marked by a high and rather emphatic nuclear accent.

No significant differences between the RefLex combinations in terms of accent type could be found. Thus, all sub-hypotheses have to be rejected, as can be seen in Table 5, which summarises our results for the spontaneous data.

6. Conclusions

The analysis of two small corpora of read and spontaneous German widely confirms the relevance of both a referential and a lexical level for an investigation of information status in spoken language. Crucially, *all* hypotheses on the distribution of accent *positions* could be confirmed for read speech, which shows a general tendency of a stepwise increase in prosodic prominence from given to new items. It also becomes obvious that an item's information status *at each level separately* has an incremental effect on the degree of prosodic prominence assigned to the item. That is, referentially and lexically new expressions are most prominent, and referentially and lexically given expressions are least prominent, whereas this difference is levelled out in combinations of givenness and newness (i.e. in r-new/l-given vs. r-given/l-new). This picture is less clear in spontaneous speech, however. Nevertheless, a distinction between fully

new (r-new/l-new) and fully given non-pronominal items (r-given/l-given) in the distribution of accent positions could also be found in the spontaneous data.

Our results confirm that the claims made in the literature about deaccentuation versus accentuation (e.g. Allerton, 1978; Cruttenden, 2006) are much too strong. Fewer cases of deaccentuation than expected could be found in both data sets. Instead, the data revealed gradual differences in prosodic prominence, reflecting differences in the *probability* of an item to be (de-)accented.

Several reasons are conceivable why most hypotheses could not be confirmed for the spontaneous data in the same way as for the read data, and why there is generally a higher percentage of nuclear accents for all label combinations. First of all, speakers of spontaneous speech tend to produce shorter intonation phrases, which by definition contain (at least) one accent, i.e. the nuclear accent. As a consequence, all information status combinations, i.e. new as well as given elements on various levels, are mostly marked by nuclear accents. This means that possible other accent positions implying lower levels (and in general different shades) of prominence are overridden, which also leads to a reduced variability in terms of accent types. Secondly, an analysis of the data shows more instances of predications[23] and more overtly contrastive constellations than in the controlled texts which serve as a basis for the read data. Both predications and contrastive (or alternative-eliciting) expressions have the potential to override the prosodic marking of information status. Thirdly, the spontaneous data set is much smaller than the read data set, so that some effects might have been obscured due to an insufficient amount of cases.

As far as focus/contrast is concerned, Rooth (1992) and, more recently, Selkirk (2008), Beaver and Velleman (2011), and Riester and Baumann (2013) have provided features which trigger contrastive interpretations, which in turn attract the strongest prominence in the sentence, for instance, answers to wh-questions (which play no role in our data sets since they do not contain dialogues) and association with focus-sensitive particles (prominent in the semantic literature but rare in our data). The most important of the well-known triggers which are found in our spontaneous data are overt contrasts. An example is provided in Fig. 11.

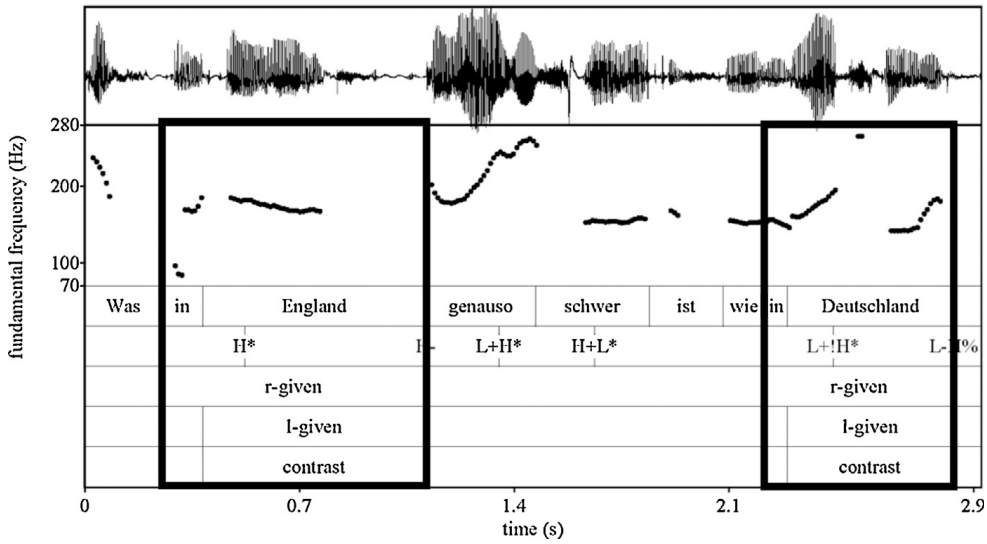

Fig. 11. Two instances of fully given but contrastive expressions: pitch contour and label tiers of the spontaneous utterance *Was in England genauso schwer ist wie in Deutschland* ('which is just as difficult in England as it is in Germany').

The two fully given expressions (*England* and *Deutschland* have been mentioned in the immediate context) are realised by means of nuclear pitch accents (one H* and one !H*). It is obvious that the more frequently such constellations occur in a text the more will they add to the degree of prominence found for given information. Such contrastive structures received a pitch accent in 79% of the cases in our spontaneous data.

[23] In predicative and identificational constructions, nominal expressions are used to assert properties of existing discourse entities. This may lead to the assignment of a nuclear accent to the nominal attribute, regardless of its information status, compare example (20).

(20) Pete is the BOSS.

It is beyond the scope of the current article to annotate predicative constructions (but see e.g. the treatment of attributive expressions in the *OntoNotes* project, cf. Weischedel et al. (2011:21).

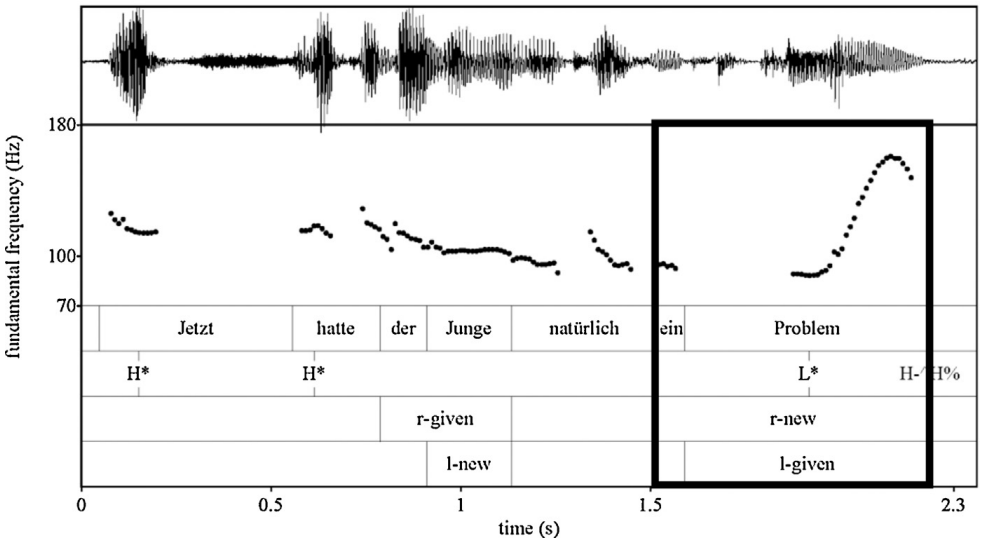

Fig. 12. Pitch contour and label tiers of the spontaneous utterance *Jetzt hatte der Junge natürlich ein Problem* ('now the boy naturally had a problem').

Another – phonological – reason for increased prominence on somewhat activated information is the rhythmic structure of utterances which may require additional accents in *prenuclear* position (see Féry and Kügler (2008) and Section 2 above). Prenuclear accents were more frequent in the read data, which generally displays longer intonation phrases than spontaneous speech.

Our overall results (in particular for read speech) suggest that accent *position* is a more decisive prosodic marker of information status than accent *type*. Moreover, it seems that the choice of accent type is to some extent *determined* by accent position, since certain accent types are more likely to occur in prenuclear or nuclear position, respectively.

Furthermore, the distribution of accent types reveal differences in the speaking style of read and spontaneous speech in German. In particular, there is a clear distinction between speakers' preferred *nuclear* accent type: whereas intonation phrases in read speech often end in Early Peak accents (29% of the cases), the nuclear pitch movement in the spontaneous data was most often characterised by the combination of a low nuclear accent and a high boundary tone (42% of the cases; cf. the example in Fig. 12). That is, in narration monologues speakers make ample use of continuation rises.[24]

Although the types of accent used in both data sets did not vary as much as expected, the weighting procedure for the read data reveals an interesting result (see Fig. 6a), clearly supporting our general hypothesis: the accent type distribution reflects a higher degree of prominence in correspondence with a higher degree of newness at the lexical level. As for prominence, the label combinations cluster along the l-level, with the r-labels figuring as secondary ordering criteria. Generally, this grouping suggests that the lexical level of an item's information status may be the predominant trigger for its prosodic realisation.

With the approach proposed in this paper we hope to have opened a perspective for a detailed inspection of information-theoretically distinct expressions in various types of spoken language. The simpler and more intuitive an annotation scheme is the easier and more straightforward it is to create annotated corpora of reasonable sizes.[25] The RefLex scheme presented here has proven appropriate in this respect and is thus a promising tool for the investigation of the relation between an item's information status and its prosodic realisation, which still has to be extended to other linguistic levels of description, in particular contrastive/alternative-eliciting information as well as certain syntactic constellations which also have an influence on prosody.

Acknowledgements

We would like to thank Jutta Hartmann, Oliver Bott as well as two anonymous reviewers for their extremely valuable comments on earlier versions of the paper. For their assistance during various steps in the analysis of the spoken data we

[24] Certainly, storytelling is a specific kind of data, whose characteristics cannot simply be transferred to other types of spontaneous speech. This particularly holds for a conversation between two or more interlocutors, which is the most natural way of human communication. Thus, more natural data in dialogue form is needed for comparison.

[25] A current enterprise is the re-annotation of the DIRNDL corpus (Eckart et al., 2012) with RefLex labels. DIRNDL is a 50.000-word corpus of German radio news also labelled for syntax and prosody.

thank Christine Röhr, Marlene Biehl, Simon Ritter and Rainer Schmitz. Financial support by the German Science Foundation (DFG) is kindly acknowledged. (Project A1 of the Stuttgart Collaborative Research Center (SFB) 732 Incremental Specification in Context, and the project GR 1610/5 Degrees of Activation and Focus-Background Structure in Spontaneous Speech in Cologne.)

References

Allerton, D.J., 1978. The notion of 'Givenness' and its relation to presupposition and theme. Lingua 44, 133–168.
Ayers, G., 1996. Nuclear accent types and prominence: some psycholinguistic experiments (Unpublished doctoral dissertation). Ohio State University, Columbus, OH.
Baumann, S., 2006. The intonation of givenness – evidence from German (Linguistische Arbeiten 508). Niemeyer, Tübingen.
Baumann, S., 2012. Types of Secondary Prominence and their Relation to Information Structure. Talk at Satellite Workshop of GLOW 35 on Production and Perception of Prosodically-Encoded Information Structure. University of Potsdam, www.ling.uni-potsdam.de/(glow/w2/baumann-handout.pdf.
Baumann, S., Grice, M., 2006. The intonation of accessibility. Journal of Pragmatics 38, 1636–1657.
Baumann, S., Riester, A., 2010. Annotating information status in spontaneous speech. In: Proceedings Speech Prosody 2010, Chicago. 100092, pp. 1–4.
Baumann, S., Riester, A., 2012. Referential and lexical givenness: semantic, prosodic and cognitive aspects. In: Elordieta, G., Prieto, P. (Eds.), Prosody and Meaning. Mouton De Gruyter, Berlin, New York. (Interface Explorations 25), pp. 119–162..
Baumann, S., Becker, J., Grice, M., Mücke, D., 2007. Tonal and articulatory marking of focus in German. In: Proceedings 16th ICPhS, Barcelona. pp. 1029–1032.
Beaver, D., Geurts, B., 2011. Presupposition. Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/entries/presupposition/
Beaver, D., Velleman, D., 2011. The communicative significance of primary and secondary accents. Lingua 121, 1671–1692.
Boersma, P., Weenink, D., 1996. PRAAT, a system for doing phonetics by computer (version 3.4.). Report 132, Institute of Phonetic Sciences of the University of Amsterdam, http://www.fon.hum.uva.nl/praat/
Büring, D., 2002. Information structure. In: Talk presented at FFL, Düsseldorf.
Büring, D., 2007. Intonation, semantics and information structure. In: Ramchand, G., Reiss, C. (Eds.), The Oxford Handbook of Linguistic Interfaces. Oxford University Press, Oxford.
Burkhardt, P., 2006. Inferential bridging relations reveal distinct neural mechanisms: evidence from event-related brain potentials. Brain and Language 98 (2), 159–168.
Burkhardt, P., 2007. The P600 reflects cost of new information in discourse memory. Neuroreport 18 (17), 1851–1854.
Burkhardt, P., Roehm, D., 2007. Differential Effects of saliency: an event-related brain potential study. Neuroscience Letters 413 (2), 115–120.
Chafe, W., 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In: Li, C. (Ed.), Subject and Topic. Academic Press, New York, pp. 25–56.
Chafe, W., 1987. Cognitive constraints on information flow. In: Tomlin, R. (Ed.), Coherence and Grounding in Discourse. John Benjamins, Amsterdam, pp. 21–52.
Chafe, W., 1994. Discourse, Consciousness, and Time. University of Chicago Press, Chicago/London.
Clark, H., 1977. Bridging. In: Johnson-Laird, P.N., Wason, P.C. (Eds.), Thinking: Readings in Cognitive Science. Cambridge University Press, Cambridge, pp. 411–420.
Clark, H., Haviland, S., 1977. Comprehension and the given-new contract. In: Freedle, R. (Ed.), Discourse Production and Comprehension. Ablex, NJ, pp. 1–40.
Clark, H., Sengul, C.J., 1979. In search of referents for nouns and pronouns. Memory and Cognition 7, 35–41.
Cruttenden, A., 2006. The de-accenting of given information: a cognitive universal? In: Bernini, G., Schwartz, M. (Eds.), Pragmatic Organization of Discourse in the Languages of Europe. Mouton de Gruyter, The Hague, pp. 311–356.
Crystal, D., 1969. Prosodic Systems and Intonation in English. Cambridge University Press, Cambridge.
Eckart, K., Riester, A., Schweitzer, K., 2012. A discourse information radio news database for linguistic analysis. In: Chiarcos, C., Nordhoff, S., Hellmann, S. (Eds.), Representing Language Data and Metadata. Springer, Heidelberg, pp. 65–76.
Eckert, M., Strube, M., 2000. Dialogue acts, synchronizing units and anaphora resolution. Journal of Semantics 17 (1), 51–89.
Féry, C., Kügler, F., 2008. Pitch accent scaling on given, new and focused constituents in German. Journal of Phonetics 36, 680–703.
Götze, M., Endriss, C., Hinterwimmer, S., Fiedler, I., Petrova, S., Schwarz, A., Skopeteas, S., Stoel, R., Weskott, T., 2007. Information structure. In: Dipper, S., Götze, M., Skopeteas, S. (Eds.), Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics, and Information Structure. Working Papers of the CRC 632, Interdisciplinary Studies on Information Structure (ISIS) 7.) Universitätsverlag Potsdam, Potsdam, pp. 147–187.
Grice, M., Ladd, D.R., Arvaniti, A., 2000. On the place of phrase accents in intonational phonology. Phonology 17 (2), 143–185.
Grice, M., Baumann, S., Benzmüller, R., 2005. German intonation in autosegmental-metrical phonology. In: Jun, S. (Ed.), Prosodic Typology: The Phonology of Intonation and Phrasing. Oxford University Press, Oxford, pp. 55–83.
Grice, M., Baumann, S., Jagdfeld, N., 2009. Tonal association and derived nuclear accents: the case of downstepping contours in German. Lingua 119, 881–905.
Gundel, J., Hedberg, N., Zacharski, R., 1993. Cognitive status and the form of referring expressions in discourse. Language 69, 274–307.
Gussenhoven, C., 2004. The Phonology of Tone and Intonation. Cambridge University Press, Cambridge.
Halliday, M.A.K., 1967. Notes on transitivity and theme in English, part 2. Journal of Linguistics 3, 199–244.
Halliday, M.A.K., Hasan, R., 1976. Cohesion in English. Longman, London.
Halliday, M.A.K., Matthiessen, C., 2004. An Introduction to Functional Grammar, 3rd rev. ed. Edward Arnold, London.
Heim, I., 1982. On the Semantics and Indefinite Noun Phrases. University of Massachusetts at Amherst, (Ph.D. thesis).

Jagdfeld, N., Baumann, S., 2011. Order effects on the perception of relative prominence. In: Proceedings 17th ICPhS, Hongkong, China, pp. 958–961.
Kohler, K.J., 1991. Terminal intonation patterns in single-accent utterances of German: phonetics, phonology, and semantics. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel 25, 115–185.
Krifka, M., 2007. Basic notions of information structure. In: Féry, C., Fanselow, G., Krifka, M. (Eds.), Working Papers of the CRC 632, Interdisciplinary Studies on Information Structure (ISIS) 6. Universitätsverlag Potsdam, Potsdam, pp. 13–55.
Ladd, D.R., 1980. The Structure of Intonational Meaning: Evidence from English. Indiana University Press, Bloomington.
Ladd, D.R., 2008. Intonational Phonology, second ed. Cambridge University Press, Cambridge.
Ladd, D.R., Morton, R., 1997. The perception of intonational emphasis: continuous or categorical? Journal of Phonetics 25, 313–342.
Lambrecht, K., 1994. Information Structure and Sentence Form. Cambridge University Press, Cambridge.
Liberman, M., 1975. The Intonational System of English. Garland, New York.
Löbner, S., 1998. Definite associative anaphora. In: Botley, S. (Ed.), Approaches to Discourse Anaphora. Proceedings of the Discourse Anaphora and Resolution Colloquium. Lancaster University, Lancaster.
Nissim, M., Dingare, S., Carletta, J., Steedman, M., 2004. An annotation scheme for information status in dialogue. In: Proceedings of the 4th LREC, Lisbon. pp. 1023–1026.
Pierrehumbert, J., Hirschberg, J., 1990. The meaning of intonational contours in the interpretation of discourse. In: Cohen, P.R., Morgan, J., Pollack, M.E. (Eds.), Intentions in Communication. MIT Press, Cambridge, MA, pp. 271–311.
Prince, E.F., 1981. Toward a taxonomy of given-new information. In: Cole, P. (Ed.), Radical Pragmatics. Academic Press, NY, pp. 223–256.
Prince, E.F., 1992. The ZPG letter: subjects, definiteness, and information-status. In: Thompson, S., Mann, W. (Eds.), Discourse Description: Diverse Analyses of a Fund Raising Text. John Benjamins, Amsterdam, pp. 295–325.
Riester, A., 2008. A semantic explication of 'Information Status' and the underspecification of the recipients' knowledge. In: Grønn, A. (Ed.), Proceedings of Sinn und Bedeutung 12. University of Oslo, Oslo, pp. 508–522.
Riester, A., Baumann, S., 2013. Focus triggers and focus types from a corpus perspective. In: Dipper, S., Zinsmeister, H., Webber, B. (Eds.), Dialogue and Discourse 4(2) (Special Issue 'Annotating Pragmatic and Discourse Phenomena'), pp. 215–248. http://dx.doi.org/10.5087/dad.2013.210.
Riester, A., Lorenz, D., Seemann, N., 2010. A recursive annotation scheme for referential information status. In: Proceedings of the Seventh International Conference of Language Resources and Evaluation (LREC), Valletta, Malta, pp. 717–722.
Röhr, C., Baumann, S., 2010. Prosodic marking of information status in German. In: Proceedings Speech Prosody 2010, Chicago. 100019, pp. 1–4.
Röhr, C., Baumann, S., 2011. Decoding information status by type and position of accent in German. In: Proceedings 17th ICPhS, Hong Kong, China, pp. 1706–1709.
Rooth, M., 1992. A theory of focus interpretation. Natural Language Semantics 1 (1), 75–116.
Schumacher, P., Baumann, S., 2010. Pitch accent type affects the N400 during referential processing. NeuroReport 21 (9), 618–622.
Schwarzschild, R., 1999. GIVENness, AvoidF, and other constraints on the placement of accent. Natural Language Semantics 7 (2), 141–177.
Selkirk, E., 2008. Contrastive focus, givenness and the unmarked status of 'Discourse-New'. In: Féry, C., Fanselow, G. (Eds.). Acta Hungarica Linguistica 55 (3–4), 331–346, Special Issue on Information Structure.
Stalnaker, R., 1974. Pragmatic presuppositions. In: Munitz, M., Unger, P. (Eds.), Semantics and Philosophy. New York University Press, New York, pp. 197–214.
Uhmann, S., 1991. Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie. Niemeyer, Tübingen.
van Deemter, K., 1994. What's new? A semantic perspective on sentence accent. Journal of Semantics 11, 1–31.
van Deemter, K., 1999. Contrastive stress, contrariety, and focus. In: Bosch, P., van der Sandt, R. (Eds.), Focus – Linguistic, Cognitive, and Computational Perspectives (Studies in Natural Language Processing). Cambridge University Press, Cambridge, pp. 3–17.
van der Sandt, R., 1992. Presupposition projection as anaphora resolution. Journal of Semantics 9, 333–377.
Vieira, R., Poesio, M., 2000. Processing definite descriptions in corpora. In: Botley, S., McEnery, T. (Eds.), Corpus-based and Computational Approaches to Anaphora. UCL Press, London.
Weischedel, R., Palmer, M., Marcus, M., Hovy, E., Pradhan, S., Ramshaw, L., Xue, N., Taylor, A., Kaufman, J., Franchini, M., El-Bachouti, M., Belvin, R., Houston, A., 2011. OntoNotes Release 4. 0. Linguistic Data Consortium, Philadelphia.